



To: Steve Bamberg, Ed Choroski, Rick Gates, Tim Gorman, Scott Jonkman,
Brian Leander, Barb Martin, Gary Peters, Dwayne Pinon, Linda Rambo,
James VanOverbake

Cc: Joseph Tiemeyer

From: Wayne Bancroft
Tracy Morris

Regarding: DEA Suspicious Order Reporting
Date: June 23, 2008

Deliverable: Proposal for defining “suspicious orders” in the Walgreen distribution
system

Overview

The DEA is requiring Walgreens to monitor the orders, for control substances, that our stores place on our distribution centers for suspicious activity. Suspicious orders are defined in terms of order size and order frequency. This document proposes a methodology for identifying suspicious orders in terms of order size and order frequency. First the reasoning behind the method is described. Followed by the steps needed to perform the analysis. The Appendix gives examples using real data.

To monitor the **orders size**, tolerance limits will be established for each store/item combination. If an order is placed on the DC that exceeds its tolerance limit the order is flagged as suspicious.

To monitor **orders frequency** the geometric distribution can be used to determine the probability $P(n)$ that it will be n weeks or more till the next order is created. Using $P(n)$ we determine the minimum number till the next order deemed acceptable. If the next orders is placed earlier than expected the order is flagged as suspicious.

In either case, if the order quantity does not exceed the SIMS suggested order quantity then the order should no longer be considered suspicious. If the order is identified as suspicious, a detailed report should be created to aid the analysis that has to make a quick decision to allow or stop the order. See Appendix B.

Tolerance Limits

To monitor the orders size (Q), tolerance limits are established for each item. It is assumed that the order sizes are normally distributed with a fixed mean and standard deviation. Further more, the frequencies of orders are assumed to fluctuate in accordance with the customer demand patterns. In this way, the application is applicable to any type of demand patterns. In other words, the customer demands may not be normally distributed, however the size of the orders placed on the DC will be normally distributed.

The following steps are followed to establish the order size upper tolerance limit: (Tolerance Limits should be reestablished periodically – i.e. monthly.)

- 1) Using all the orders for the past T weeks (i.e. 52), find the sample mean (\bar{Q}) and the standard deviation (S_Q).
- 2) Let N be the total number of orders placed in time period T .
- 3) With N observations table 1 is used to find the appropriate value of K for given values of P_1 and P_2 . (See explanation of P_1 and P_2 that follows.)
- 4) The upper limit for the order size is found for each item.
- 5) If an order is placed on the DC for an item and the order quantity exceeds the upper limit it is flagged as suspicious.
- 6) If the order was generated by SIMS and the order size \leq suggested order size it is no longer classified as suspicious. (If an order is flagged as suspicious it should not be used when tolerance limits are reestablished for an item.

Note: The item is not a candidate if the order quantity $Q = 0$ or $S_Q = 0$ or $N \leq 2$.

Using tolerance limits a value of K is found so that one can assert with confidence P_1 that the proportion of orders sizes $Q \leq \bar{Q} + K \cdot S_Q$ is at least P_2 . Note: \bar{Q} is the mean order quantity and S_Q is the standard deviation estimated from the data. Table 1 lists K value for $P_1 = (0.95, 0.99)$ and $P_2 = (0.90, 0.95, 0.99)$ and for selected values of sample size N . For example, with a sample size of 40, $P_1 = 0.99$ and $P_2 = 0.95$, $K =$

3.21 and the upper tolerance limit becomes $\bar{Q} + 3.21 \cdot S_Q$. (This limit should be round to the nearest

multiple.) This means that with a 99% confidence, the proportion of the population within ($\bar{Q} + 3.21 \cdot S_Q$) is at least 0.95. Any entry beyond that limit is questionable.

Order Quantity Logic Matrix:

		Order Qty vs Tolerance Limit	
		$Q > TL$	$Q \leq TL$
Order Qty vs Suggested Qty	$Q > S_Q$	Flagged as Suspicious	OK
	$Q \leq S_Q$	OK	OK

Order Frequency

If two consecutive orders are generated relatively close together in comparison to past history it may be considered suspicious. We can use the geometric distribution to establish the minimum number of weeks¹ (min_weeks) until the next orders deemed acceptable. If an order is created earlier than the min_weeks it is flagged as suspicious.

(Unlike the tolerance limits we are assuming the demand for an item is normally distributed. Thus, seasonal items are not good candidates for this application. We can either recalibrate the mode in and out of season or skip seasonal items all together.)

The following steps are followed to establish the minimum acceptable number of week between orders:

- 1) Estimate the probability P an order will be generated in a given week. Let N be the total number of week of order history used to determine P (i.e. $N = 52$). Next let n be the number of weeks the item was order from the DC. Then the estimate of P becomes $P = n / N$.
- 2) The probability that an order will not be placed in a week becomes $Q = (1 - P)$
- 3) Next we want to determine the probability $P(n)$ that the next order will be placed in week n :

$$P(n) = P * Q^{(n-1)}$$
- 4) Find $F(n)$, the cumulative probability that the first order will occur in week $w \leq n$. $F(n) = \sum_{t=1}^n P(t)$ for $w = 1$ to n .
- 5) Find the minimum value of n where $F(n) \geq (1 - P')$. Where we are confident that at least P' (i.e. 98%) of the orders will be placed beyond the n th week. Thus, less than $1 - P'$ (<2%) of the orders will occur within n week of the last order.
- 6) Don't allow n to exceed 13 weeks. This would represent a 90-day supply.

Order Frequency Logic Matrix:

w = next order in weeks		Order Frequency	
n = min #weeks acceptable		w < n	w >= n
Order Qty vs Suggested Qty	Q > SQ	Flagged as Suspicious	OK
	Q <= SQ	OK	OK

¹ Alternative, days can be used in place of weeks

Appendix A

Table 1: Values of K for Tolerance Limits for Normal Distributions.

n\P ₂	P ₁ = 0.95			P ₁ = 0.99		
	0.9	0.95	0.99	0.9	0.95	0.99
2	32.019	37.674	48.43	160.193	188.491	242.3
3	8.38	9.916	12.861	18.93	22.401	29.055
4	5.369	6.37	8.299	9.398	11.15	14.527
5	4.275	5.079	6.634	6.612	7.855	10.26
6	3.712	4.414	5.775	5.337	6.345	8.301
7	3.369	4.007	5.248	4.613	5.488	7.187
8	3.136	3.732	4.891	4.147	4.936	6.468
9	2.967	3.532	4.631	3.822	4.55	5.966
10	2.839	3.379	4.433	3.582	4.265	5.594
11	2.737	3.259	4.277	3.397	4.045	5.308
12	2.655	3.162	4.15	3.25	3.87	5.079
13	2.587	3.081	4.044	3.13	3.727	4.893
14	2.529	3.012	3.955	3.029	3.608	4.737
15	2.48	2.954	3.878	2.945	3.507	4.605
16	2.437	2.903	3.812	2.872	3.421	4.492
17	2.4	2.858	3.754	2.808	3.345	4.393
18	2.366	2.819	3.702	2.753	3.279	4.307
19	2.337	2.784	3.656	2.703	3.221	4.23
20	2.31	2.752	3.615	2.659	3.168	4.161
25	2.208	2.631	3.457	2.494	2.972	3.904
30	2.14	2.549	3.35	2.385	2.841	3.733
35	2.09	2.49	3.272	2.306	2.748	3.611
40	2.052	2.445	3.213	2.247	2.677	3.518
45	2.021	2.408	3.165	2.2	2.621	3.444
50	1.996	2.379	3.126	2.162	2.576	3.385
55	1.976	2.354	3.094	2.13	2.538	3.335
60	1.958	2.333	3.066	2.103	2.506	3.293
65	1.943	2.315	3.042	2.08	2.478	3.257
70	1.929	2.299	3.021	2.06	2.454	3.225
75	1.917	2.285	3.002	2.042	2.433	3.197
80	1.907	2.272	2.986	2.026	2.414	3.173
85	1.897	2.261	2.971	2.012	2.397	3.15
90	1.889	2.251	2.958	1.999	2.382	3.13
95	1.881	2.241	2.945	1.987	2.368	3.112
100	1.874	2.233	2.934	1.977	2.355	3.096
150	1.825	2.175	2.859	1.905	2.27	2.983
200	1.798	2.143	2.816	1.865	2.222	2.921
250	1.78	2.121	2.788	1.839	2.191	2.88
300	1.767	2.106	2.767	1.82	2.169	2.85
400	1.749	2.084	2.739	1.794	2.138	2.809
500	1.737	2.07	2.721	1.777	2.117	2.783
600	1.729	2.06	2.707	1.764	2.102	2.763
700	1.722	2.052	2.697	1.755	2.091	2.748
800	1.717	2.046	2.688	1.747	2.082	2.736
900	1.712	2.04	2.682	1.741	2.075	2.726
1000	1.709	2.036	2.676	1.736	2.068	2.718
∞	1.645	1.96	2.576	1.645	1.96	2.576

Examples: Tolerance Limits

Tolerance Limits									
Store:	From:	1071224	To:	1080618	With Rounded Upper Limit				
PLN	Avg Q	STD of Q	Number of Orders	K value *	Upper Limit	Max Order Qty	Outlier (y/n)	Suggested Order Qty	Outlier (y/n)
40000553697	2.5	1.1	45	2.68	5	6	y	6	n
40000563756	1.1	0.3	11	4.05	2	2	n	2	n
40000672664	1.4	0.6	38	2.75	3	3	n	3	n
40000672713	1.1	0.4	26	2.94	2	3	y	3	n
40000673037	3.2	2.0	39	2.69	9	9	n	9	n
40000673665	2.7	2.5	20	3.17	10	11	y	11	n
40000674231	1.2	0.4	32	2.84	2	3	y	3	n
40000675756	1.5	0.7	39	2.69	3	4	y	3	y
Total							5		1
* P ₁ = 99% and P ₂ = 95%									
Tolerance Limits									
Store:	From:	1071224	To:	1080618	With Rounded Upper Limit				
PLN	Avg Q	STD of Q	Number of Orders	K value *	Upper Limit	Max Order Qty	Outlier (y/n)	Suggested Order Qty	Outlier (y/n)
40000553697	1.9	0.9	41	2.68	4	5	y	4	n
40000560610	3.6	2.3	37	2.75	10	12	y	12	y
40000673036	1.0	0.0	8	4.94	1	1	n	1	n
40000674484	1.5	1.0	22	3.08	4	5	y	5	n
40000675041	1.5	0.9	31	2.84	4	5	y	4	y
40000684278	1.1	0.4	8	4.94	3	2	n	1	n
Total							4		2
Tolerance Limits									
Store:	From:	1071224	To:	1080618	With Rounded Upper Limit				
PLN	Avg Q	STD of Q	Number of Orders	K value *	Upper Limit	Max Order Qty	Outlier (y/n)	Suggested Order Qty	Outlier (y/n)
40000563756	1.1	0.3	18	3.279	2	2	n	2	n
40000674484	1.7	1.0	34	2.841	4	5	y	5	n
Total							1		0
Tolerance Limits									
Store:	From:	1071224	To:	1080618	With Rounded Upper Limit				
PLN	Avg Q	STD of Q	Number of Orders	K value *	Upper Limit	Max Order Qty	Outlier (y/n)	Suggested Order Qty	Outlier (y/n)
40000563756	5.6	3.3	51	2.576	14	16	y	12	y
40000672663	5.6	1.7	50	2.576	10	13	y	13	n
40000675041	1.6	1.0	31	2.841	4	5	y	5	n
40000682432	1.4	0.7	43	2.677	3	4	y	4	n
40000682586	9.0	8.1	40	2.677	31	35	y	27	y
Total							5		2

The 4 store/examples listed the PLNs where the historical order quantity exceeded its tolerance limit. Overall 15 suspicious orders were identified from 912 SKUs over a 26 weeks period. Only 5 outliers remained after removing order where the order quantity <= suggested order quantity. This roughly translates into 3 items per store per year being identified as suspicious. Of the final 5 suspicious items, 2 had recent theft casts. (See below.)

The 9 items in bold had recent theft cases. (Store are not identified. See page 7 for item descriptions.) 3 of these 9 items were flagged per the tolerance limits and 2 of these items also had an order quantity > the suggested order quantity.

Example: Order Frequency

Order Frequency					F(n) : Cumulative probability that the next order will be placed in week w <= n						
PLN	Count	P	Q	Week n	1	2	3	4	5	6	7
30009041702	1	0.04	0.96	1	0.04	0.07	0.11	0.14	0.17	0.20	0.23
30071101868	6	0.22	0.78	1	0.22	0.40	0.53	0.63	0.72	0.78	0.83
30071101968	4	0.15	0.85	1	0.15	0.27	0.38	0.47	0.55	0.62	0.67
30091332201	1	0.04	0.96	1	0.04	0.07	0.11	0.14	0.17	0.20	0.23
30187049802	1	0.04	0.96	1	0.04	0.07	0.11	0.14	0.17	0.20	0.23
30525001601	1	0.04	0.96	1	0.04	0.07	0.11	0.14	0.17	0.20	0.23
30603516521	1	0.04	0.96	1	0.04	0.07	0.11	0.14	0.17	0.20	0.23
31012230110	4	0.15	0.85	1	0.15	0.27	0.38	0.47	0.55	0.62	0.67
32483922216	4	0.15	0.85	1	0.15	0.27	0.38	0.47	0.55	0.62	0.67
32483922816	1	0.04	0.96	1	0.04	0.07	0.11	0.14	0.17	0.20	0.23
35099160316	5	0.19	0.81	1	0.19	0.34	0.46	0.56	0.64	0.71	0.76
35167433004	1	0.04	0.96	1	0.04	0.07	0.11	0.14	0.17	0.20	0.23
35970279916	2	0.07	0.93	1	0.07	0.14	0.21	0.26	0.32	0.37	0.42
36437661131	3	0.11	0.89	1	0.11	0.21	0.30	0.38	0.45	0.51	0.56
36647951210	1	0.04	0.96	1	0.04	0.07	0.11	0.14	0.17	0.20	0.23
36649065020	1	0.04	0.96	1	0.04	0.07	0.11	0.14	0.17	0.20	0.23
40000522239	3	0.11	0.89	1	0.11	0.21	0.30	0.38	0.45	0.51	0.56
40000522462	19	0.70	0.30	1	0.70	0.91	0.97	0.99	1.00	1.00	1.00
40000523789	2	0.07	0.93	1	0.07	0.14	0.21	0.26	0.32	0.37	0.42
40000524280	3	0.11	0.89	1	0.11	0.21	0.30	0.38	0.45	0.51	0.56
40000532532	4	0.15	0.85	1	0.15	0.27	0.38	0.47	0.55	0.62	0.67
40000532538	12	0.44	0.56	1	0.44	0.69	0.83	0.90	0.95	0.97	0.98

This example list F(n) the cumulative probability that the next order will occur in week w <= n. The column where F(n) > 0.02 is highlighted and the associated value of n is listed in column 'Week n'. In the 4 examples all the PLNs have at least a 4% of orders being placed in two consecutive weeks. Thus, using a 98% confidence limit, no items have been identified as suspicious using. The probability estimates were based on 26 weeks of data.

Note: The frequency model could also converted to a daily (order creation day) model as follows:

Order Frequency				Limit	Actual	F(n) : Cumulative probability that the next order will be placed in day d <= n						
PLN	Count	P	Q	Day n	min Days	1	2	3	4	5	6	7
30009041702	1	0.01	0.99	4	na	0.006	0.011	0.017	0.022	0.028	0.033	0.039
30071101868	6	0.03	0.97	1	8	0.034	0.066	0.098	0.128	0.158	0.186	0.213
30071101968	4	0.02	0.98	1	7	0.022	0.044	0.066	0.087	0.107	0.127	0.147
30091332201	1	0.01	0.99	2	3	0.011	0.022	0.033	0.044	0.055	0.066	0.076
30187049802	1	0.01	0.99	4	na	0.006	0.011	0.017	0.022	0.028	0.033	0.039
30525001601	1	0.01	0.99	4	na	0.006	0.011	0.017	0.022	0.028	0.033	0.039
30603516521	1	0.01	0.99	4	na	0.006	0.011	0.017	0.022	0.028	0.033	0.039
31012230110	4	0.02	0.98	1	11	0.022	0.044	0.066	0.087	0.107	0.127	0.147
32483922216	4	0.02	0.98	1	7	0.022	0.044	0.066	0.087	0.107	0.127	0.147
32483922816	1	0.01	0.99	4	na	0.006	0.011	0.017	0.022	0.028	0.033	0.039
35099160316	5	0.03	0.97	1	6	0.028	0.055	0.082	0.108	0.133	0.157	0.181
35167433004	1	0.01	0.99	2	1	0.011	0.022	0.033	0.044	0.055	0.066	0.076
35970279916	2	0.01	0.99	2	28	0.011	0.022	0.033	0.044	0.055	0.066	0.076
36437661131	3	0.02	0.98	2	14	0.017	0.033	0.050	0.066	0.081	0.097	0.112
36647951210	1	0.01	0.99	4	na	0.006	0.011	0.017	0.022	0.028	0.033	0.039
36649065020	1	0.01	0.99	4	na	0.006	0.011	0.017	0.022	0.028	0.033	0.039
40000522239	3	0.02	0.98	2	7	0.017	0.033	0.050	0.066	0.081	0.097	0.112
40000522462	19	0.11	0.89	1	3	0.129	0.242	0.340	0.425	0.499	0.564	0.620
40000523789	2	0.01	0.99	2	na	0.011	0.022	0.033	0.044	0.055	0.066	0.076
40000524280	3	0.02	0.98	1	2	0.022	0.044	0.066	0.087	0.107	0.127	0.147
40000532532	4	0.02	0.98	1	14	0.022	0.044	0.066	0.087	0.107	0.127	0.147
40000532538	12	0.07	0.93	1	7	0.067	0.130	0.189	0.244	0.295	0.342	0.386
na: only 1 order or Actual > 99 days												

Tolerance Limit Examples: The 9 items that had recent theft cases

PLN	Description
40000563756	HYDROCOD-APAP 10-325 TAB(WAT)+500
40000672664	HYDROCODONE/APAP 10/650 TAB +100
40000673037	ALPRAZOLAM 2MG TAB (PPC) + 100
40000675756	HYDROCOD-APAP 10-500 TB (WT)+500
40000673036	ALPRAZOLAM 1MG TAB (PPC) + 500
40000684278	OXYCONTIN 40MG TAB (PUR) 100
40000563756	HYDROCOD-APAP 10-325 TAB(WAT)+500
40000563756	HYDROCOD-APAP 10-325 TAB(WAT)+500
40000672663	HYDROCODONE/APAP 5/500+TAB(WT)500